

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/501,227

Source: PG110

Date Processed by STIC: 8/18/05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 08/18/2005

PATENT APPLICATION: US/10/501,227

TIME: 16:46:39

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw

```

3 <110> APPLICANT: SAITO, Koichi
5 <120> TITLE OF INVENTION: Estrogen receptor genes and utilization thereof
7 <130> FILE REFERENCE: 600630-21US (561334)
9 <140> CURRENT APPLICATION NUMBER: US 10/501,227
10 <141> CURRENT FILING DATE: 2004-07-12
12 <150> PRIOR APPLICATION NUMBER: JP 2002-004395
13 <151> PRIOR FILING DATE: 2002-01-11
15 <160> NUMBER OF SEQ ID NOS: 48
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 506
19 <212> TYPE: PRT
20 <213> ORGANISM: Blue Gill
22 <400> SEQUENCE: 1
23 Met Ser Leu Lys Asp Trp Leu Leu Gly Lys Glu Arg Thr Val Val Thr
24   1           5           10           15
25 Met Glu Glu Leu Arg Ser Ser Val Pro Ser Ser Gln Gln Pro Val Pro
26           20           25           30
27 Arg Glu Asp Gln Cys Ala Thr Ser Asp Glu Ser Tyr Ser Val Gly Glu
28           35           40           45
29 Ser Gly Ala Gly Ala Arg Gly Phe Glu Met Ala Lys Glu Met Arg Phe
30           50           55           60
31 Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp
32           65           70           75           80
33 Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His
34           85           90           95
35 Asn Asp Tyr Met Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Arg Asn
36           100          105          110
37 Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val
38           115          120          125
39 Gly Met Met Lys Gly Gly Val Arg Lys Asp Arg Gly Arg Val Leu Arg
40           130          135          140
41 Arg Asp Lys Arg Arg Ala Gly Thr Asn Asp Arg Glu Lys Ala Ser Lys
42           145          150          155          160
43 Asp Leu Glu Tyr Lys Thr Val Pro Pro Gln Asp Arg Arg Lys His Ser
44           165          170          175
45 Ser Ser Ser Ser Ala Gly Gly Gly Gly Gly Lys Ser Ser Val Thr Gly
46           180          185          190
47 Met Ser Pro Asp Gln Val Leu Leu Leu Leu Gln Gly Ala Glu Pro Pro
48           195          200          205
49 Met Leu Cys Ser Arg Gln Lys Leu Ser Arg Pro Tyr Thr Glu Val Thr
50           210          215          220
51 Ile Met Thr Leu Leu Thr Ser Met Ala Asp Lys Glu Leu Val His Met
52           225          230          235          240

```

RAW SEQUENCE LISTING

DATE: 08/18/2005

PATENT APPLICATION: US/10/501,227

TIME: 16:46:39

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw

```

53 Ile Thr Trp Ala Lys Lys Leu Pro Gly Phe Leu Gln Leu Ser Leu His
54           245           250           255
55 Asp Gln Val Gln Leu Leu Glu Ser Ser Trp Leu Glu Val Leu Met Ile
56           260           265           270
57 Gly Leu Ile Trp Arg Ser Ile His Cys Pro Gly Lys Leu Ile Phe Ala
58           275           280           285
59 Gln Asp Leu Ile Leu Asp Arg Asn Glu Gly Asp Cys Val Glu Gly Phe
60           290           295           300
61 Val Glu Ile Phe Asp Met Leu Leu Ala Thr Ala Ser Arg Phe Arg Met
62 305           310           315           320
63 Leu Lys Leu Lys Pro Glu Glu Phe Val Cys Leu Lys Ala Ile Ile Leu
64           325           330           335
65 Leu Asn Ser Gly Ala Phe Ser Phe Cys Thr Gly Thr Met Glu Pro Leu
66           340           345           350
67 His Asn Ser Met Ala Val Gln Asn Met Leu Asp Thr Ile Thr Asp Ala
68           355           360           365
69 Leu Ile His His Ile Ser Gln Ser Gly Cys Ser Ala Gln Gln Gln Ser
70           370           375           380
71 Arg Arg Gln Ala Gln Leu Leu Leu Leu Ser His Ile Arg His Met
72 385           390           395           400
73 Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met Lys Cys Lys Asn Lys
74           405           410           415
75 Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu Asp Ala His Arg Ile
76           420           425           430
77 His Arg Pro Asp Arg Pro Ala Gln Phe Trp Ser Gln Ala Asp Gly Glu
78           435           440           445
79 Pro Pro Phe Ile Asn Asn Asn Asn Ser Ser Asn Ser Gly Ser Asn Gly
80           450           455           460
81 Gly Val Ser Ser Ser Val Gly Ser Ser Ser Gly Pro Arg Val Asn His
82 465           470           475           480
83 Glu Ser Pro Ser Arg Gly Pro Thr Gly Pro Gly Val Leu Gln Tyr Gly
84           485           490           495
85 Gly Ser Arg Ser Asp Cys Thr His Ile Leu
86           500           505
88 <210> SEQ ID NO: 2
89 <211> LENGTH: 3499
90 <212> TYPE: DNA
91 <213> ORGANISM: Blue Gill
93 <220> FEATURE:
94 <221> NAME/KEY: CDS
95 <222> LOCATION: (424)...(1944)
97 <400> SEQUENCE: 2
98 caggcagagc ccagcgcaga gcagacagcc ttgtggaaca gtactcagac ccaggatcag      60
99 ctcagccttc acagagctgg agaccctctc cccacaacgt ccctgcctc cgctgcgtgc      120
100 ccctctcagt gacatgtacc ctgaagagag cagggggtcc ggaggggtag ccactgtgga      180
101 ctttctggaa gggacctacg attatgccgc cccaccctc gccccgactc ctctttacag      240
102 ccagtctggc tactactctg tacctctgga cgcccaagg ccaccctcag atggcagcct      300
103 tcagtccctg ggcagcgggc ctaccagtcc tcttgtgttt gtgccgtcca gccccagact      360
104 cagccccttt atgcaccgcg ccagccacca ctatctggaa accacctcaa caccctcta      420

```

RAW SEQUENCE LISTING

DATE: 08/18/2005

PATENT APPLICATION: US/10/501,227

TIME: 16:46:39

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw

105	cag	atg	agt	ctg	aaa	gac	tgg	tta	tta	gga	aaa	gaa	agg	acg	gtg	gtg	468
106		Met	Ser	Leu	Lys	Asp	Trp	Leu	Leu	Gly	Lys	Glu	Arg	Thr	Val	Val	
107		1				5					10					15	
108	acc	atg	gag	gag	ctg	agg	tct	agt	gtc	cca	tcc	agc	cag	cag	cca	gtt	516
109	Thr	Met	Glu	Glu	Leu	Arg	Ser	Ser	Val	Pro	Ser	Ser	Gln	Gln	Pro	Val	
110					20					25					30		
111	ccc	aga	gag	gac	cag	tgt	gcc	acc	agt	gat	gag	tcc	tat	agt	gtg	ggg	564
112	Pro	Arg	Glu	Asp	Gln	Cys	Ala	Thr	Ser	Asp	Glu	Ser	Tyr	Ser	Val	Gly	
113				35					40					45			
114	gag	tca	ggg	gct	gga	gcc	agg	ggg	ttt	gag	atg	gcc	aag	gag	atg	cgt	612
115	Glu	Ser	Gly	Ala	Gly	Ala	Arg	Gly	Phe	Glu	Met	Ala	Lys	Glu	Met	Arg	
116			50					55					60				
117	ttc	tgt	gct	gtg	tgc	agt	gac	tat	gcc	tct	ggg	tac	cac	tac	ggg	gtg	660
118	Phe	Cys	Ala	Val	Cys	Ser	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	
119		65					70					75					
120	tgg	tcc	tgt	gaa	ggc	tgt	aag	gcc	ttc	ttt	aag	agg	agc	atc	cag	ggg	708
121	Trp	Ser	Cys	Glu	Gly	Cys	Lys	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	
122	80					85					90				95		
123	cac	aat	gac	tat	atg	tgc	cca	gca	acc	aat	cag	tgt	act	att	gac	agg	756
124	His	Asn	Asp	Tyr	Met	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp	Arg	
125					100					105					110		
126	aat	cgg	aga	aag	agc	tgc	cag	gct	tgc	cgt	ctt	agg	aag	tgt	tat	gaa	804
127	Asn	Arg	Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	
128				115					120					125			
129	gtg	ggc	atg	atg	aaa	gga	ggg	gtt	cgc	aag	gac	cgt	ggc	cgt	gtt	ttg	852
130	Val	Gly	Met	Met	Lys	Gly	Gly	Val	Arg	Lys	Asp	Arg	Gly	Arg	Val	Leu	
131			130					135					140				
132	cgc	cgt	gat	aaa	cga	cgt	gct	gga	acc	aat	gac	cga	gag	aag	gcc	tct	900
133	Arg	Arg	Asp	Lys	Arg	Arg	Ala	Gly	Thr	Asn	Asp	Arg	Glu	Lys	Ala	Ser	
134			145				150					155					
135	aag	gac	ctg	gag	tac	aaa	aca	gtg	ccc	cct	cag	gac	agg	agg	aaa	cac	948
136	Lys	Asp	Leu	Glu	Tyr	Lys	Thr	Val	Pro	Pro	Gln	Asp	Arg	Arg	Lys	His	
137	160					165					170				175		
138	agc	agc	agc	agc	agt	gcc	ggg	ggg	gga	gga	gga	aaa	tca	tca	gtg	acc	996
139	Ser	Ser	Ser	Ser	Ser	Ala	Gly	Gly	Gly	Gly	Gly	Lys	Ser	Ser	Val	Thr	
140					180					185					190		
141	ggg	atg	tct	cct	gac	cag	gtg	ctc	ctc	ctg	ctc	cag	ggg	gcc	gag	ccc	1044
142	Gly	Met	Ser	Pro	Asp	Gln	Val	Leu	Leu	Leu	Gln	Gly	Ala	Glu	Pro		
143				195					200				205				
144	cca	atg	ctg	tgc	tcc	cgt	cag	aag	ctg	agc	cga	ccg	tac	acc	gag	gtc	1092
145	Pro	Met	Leu	Cys	Ser	Arg	Gln	Lys	Leu	Ser	Arg	Pro	Tyr	Thr	Glu	Val	
146			210					215					220				
147	acc	ata	atg	aca	cta	ctc	acc	agc	atg	gcc	gat	aag	gag	ctg	gtc	cac	1140
148	Thr	Ile	Met	Thr	Leu	Leu	Thr	Ser	Met	Ala	Asp	Lys	Glu	Leu	Val	His	
149			225				230					235					
150	atg	atc	acc	tgg	gcc	aag	aag	ctt	cca	ggg	ttc	ctg	cag	ctg	tct	ctc	1188
151	Met	Ile	Thr	Trp	Ala	Lys	Lys	Leu	Pro	Gly	Phe	Leu	Gln	Leu	Ser	Leu	
152	240					245					250				255		
153	cat	gac	cag	gtg	cag	ctg	ctg	gag	agc	tcg	tgg	ctg	gag	gtg	ctg	atg	1236

RAW SEQUENCE LISTING

DATE: 08/18/2005

PATENT APPLICATION: US/10/501,227

TIME: 16:46:39

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw

154	His	Asp	Gln	Val	Gln	Leu	Leu	Glu	Ser	Ser	Trp	Leu	Glu	Val	Leu	Met	
155					260					265					270		
156	att	ggg	ctc	ata	tgg	agg	tcc	atc	cac	tgc	ccc	ggc	aaa	ctc	atc	ttc	1284
157	Ile	Gly	Leu	Ile	Trp	Arg	Ser	Ile	His	Cys	Pro	Gly	Lys	Leu	Ile	Phe	
158				275					280					285			
159	gca	cag	gac	ctc	ata	ctg	gac	agg	aat	gaa	ggg	gac	tgt	gtg	gaa	ggc	1332
160	Ala	Gln	Asp	Leu	Ile	Leu	Asp	Arg	Asn	Glu	Gly	Asp	Cys	Val	Glu	Gly	
161			290					295					300				
162	ttt	gtt	gag	atc	ttc	gac	atg	ctg	ctg	gcc	act	gcc	tcc	cgc	ttc	cgc	1380
163	Phe	Val	Glu	Ile	Phe	Asp	Met	Leu	Leu	Ala	Thr	Ala	Ser	Arg	Phe	Arg	
164		305					310					315					
165	atg	ctc	aaa	ctc	aaa	cct	gag	gag	ttt	gtc	tgc	ctc	aaa	gct	atc	atc	1428
166	Met	Leu	Lys	Leu	Lys	Pro	Glu	Glu	Phe	Val	Cys	Leu	Lys	Ala	Ile	Ile	
167	320					325					330				335		
168	ctg	ctc	aac	tct	ggt	gcc	ttc	tct	ttc	tgc	acc	ggc	aca	atg	gag	ccc	1476
169	Leu	Leu	Asn	Ser	Gly	Ala	Phe	Ser	Phe	Cys	Thr	Gly	Thr	Met	Glu	Pro	
170				340					345					350			
171	ctc	cac	aac	agc	atg	gca	gtg	cag	aac	atg	ctg	gac	acc	atc	aca	gac	1524
172	Leu	His	Asn	Ser	Met	Ala	Val	Gln	Asn	Met	Leu	Asp	Thr	Ile	Thr	Asp	
173				355					360					365			
174	gct	ctc	ata	cat	cat	atc	agc	caa	tca	gga	tgc	tcg	gct	cag	cag	cag	1572
175	Ala	Leu	Ile	His	His	Ile	Ser	Gln	Ser	Gly	Cys	Ser	Ala	Gln	Gln	Gln	
176			370					375					380				
177	tcg	agg	cgg	cag	gcc	cag	ctg	ctg	ctc	ctg	ctc	tcc	cac	atc	agg	cac	1620
178	Ser	Arg	Arg	Gln	Ala	Gln	Leu	Leu	Leu	Leu	Leu	Ser	His	Ile	Arg	His	
179		385					390					395					
180	atg	agc	aac	aaa	ggc	atg	gag	cat	ctc	tac	agc	atg	aag	tgc	aag	aac	1668
181	Met	Ser	Asn	Lys	Gly	Met	Glu	His	Leu	Tyr	Ser	Met	Lys	Cys	Lys	Asn	
182	400					405					410				415		
183	aaa	gtg	cct	ctt	tac	gac	ctt	ctg	ctg	gag	atg	ttg	gac	gct	cac	cgt	1716
184	Lys	Val	Pro	Leu	Tyr	Asp	Leu	Leu	Leu	Glu	Met	Leu	Asp	Ala	His	Arg	
185				420						425				430			
186	ata	cac	cgc	cca	gac	aga	cca	gct	cag	ttc	tgg	tcc	cag	gct	gac	gga	1764
187	Ile	His	Arg	Pro	Asp	Arg	Pro	Ala	Gln	Phe	Trp	Ser	Gln	Ala	Asp	Gly	
188				435					440					445			
189	gag	cct	ccc	ttc	att	aac	aac	aac	aac	agc	agc	aac	agt	ggc	agc	aat	1812
190	Glu	Pro	Pro	Phe	Ile	Asn	Asn	Asn	Asn	Ser	Ser	Asn	Ser	Gly	Ser	Asn	
191			450					455					460				
192	ggc	ggc	gtc	tcc	tct	tca	gtc	ggt	tcc	agt	tca	gga	ccc	cga	gtc	aac	1860
193	Gly	Gly	Val	Ser	Ser	Ser	Val	Gly	Ser	Ser	Ser	Gly	Pro	Arg	Val	Asn	
194		465					470					475					
195	cac	gag	agc	ccg	agc	aga	gga	ccc	aca	ggt	cca	gga	gtc	ctg	cag	tac	1908
196	His	Glu	Ser	Pro	Ser	Arg	Gly	Pro	Thr	Gly	Pro	Gly	Val	Leu	Gln	Tyr	
197	480					485				490				495			
198	gga	ggg	tcc	cgc	tct	gac	tgc	acc	cac	atc	cta	tga	ggccgagcac	aacaaa			1960
199	Gly	Gly	Ser	Arg	Ser	Asp	Cys	Thr	His	Ile	Leu						
200				500					505								
201	catctgaagg	tcaaaagtaa	tttttacaga	tgatgtgtgt	tgtagagaat	gaaagctaaa											2020
202	ggttgatattt	taattaattt	catgagataa	ttattttataa	attaagtgtat	tttatagttg											2080

RAW SEQUENCE LISTING

DATE: 08/18/2005

PATENT APPLICATION: US/10/501,227

TIME: 16:46:39

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw

```

203 taactgtttt agggagtttt ttttcctttg cactaatcta gttcactaca acacgagctt 2140
204 caatgcaggc aatctactat gctgcctttc ataatatctg tgattctgag tgagtacagc 2200
205 ttaatttttc caggtgttag gtcataattgt ggcactcagc tatgggtgatt tgaaatgaca 2260
206 agcagctaata ttgcctttgt atttgccctca accaaagtgc acttcttctt gggtttattg 2320
207 ggcattgttt ttactttttac atattgggat taggatgatac agacactaaa ctatgataaa 2380
208 aaacagggttc aaatgaatgt gtgattttatt ttgtgtttaa attccaacat cattaaagag 2440
209 cctgaacgtc aggtattgtg tcttaagcgt gcacgcaaac tttaaacttc tggaaaacaa 2500
210 atattttctat gatgaaatta taaaattaac agtgattgag gatgtatgtt gaattcagag 2560
211 tagatacaat ttgcacaatc aaatcctaga gcactgatca cattatgaaa gaagcaaagc 2620
212 tttcacaact ttattgttgg gtaacttcac cacatccagc tttttgtgaa tggtaggttt 2680
213 gttctgtagg cttacatgca caagagtttt ttttctgaat ttgagatatt ttatgtgtgt 2740
214 ctgcaagaga aagactgaga aatctgagga aatttgctat aagtggccct aagctttcta 2800
215 tcttgatgca gttcagaatt tcaaaatgtt actattcatc cactaattca gtgattacat 2860
216 gttgagtttg gcttgattta cacaactcca aaagcctagt tatcattaaa tatgtgcata 2920
217 tgcaattgtt tttattttgt ttaaattggaa caaattttaat ctaaatctaa tatggacctg 2980
218 accaggtgtt ttctttatta gctgctatac actgctaagc accattgtta atagtttggt 3040
219 ttatatagct aaatagcttt ttccatgacc atcaaaggcc tccaaaagaa agctaattgtt 3100
220 ctttcctaata tctgtgataa acagactcca aaatcacact ggatgggtcac tgaacaagtc 3160
221 ctgcttcatg tttgtttaca tgtcaaccag caacgtcagc acacctgtgc tgttttgtat 3220
222 cctcccatga acagttgttc agtcacaggt ttgtcacaca ggtagaacaa tctgttaata 3280
223 tactgaaaaa aaggccagag gttgacgtgt agagaatgtt gccagaatac aaatgataaa 3340
224 caaagatctg tgccacttaa acaagaatgg aaagcctcta tacagggtca ggaaactgga 3400
225 tttggacact tgaagtgcaa gtcaaacctg acctgtcatc tgtttactgt gcataaaaaat 3460
226 aaaaacatta attgggaaaa aaaaaaaaaa aaaaaaaaaa 3499
228 <210> SEQ ID NO: 3
229 <211> LENGTH: 996
230 <212> TYPE: DNA
231 <213> ORGANISM: Blue Gill
233 <400> SEQUENCE: 3
234 aggagcatcc aaggtcacaa tgactacatg tgcccagcaa ccaatcagtg tactattgac 60
235 aggaatcgga gaaagagctg ccaggcttgc cgtcttagga agtgttatga agtgggcatg 120
236 atgaaaggag gtgttcgcaa ggaccgtggc cgtgttttgc gccgtgataa acgacgtgct 180
237 ggaaccaatg accgagagaa ggctctctaa gacctggagt acaaaacagt gccccctcag 240
238 gacaggagga aacacagcag cagcagcagt gccggtggtg gaggaggaaa atcatcagtg 300
239 accgggatgt ctctgacca ggtgctcctc ctgctccagg gtgccgagcc ccaatgctg 360
240 tgctcccgtc agaagctgag ccgaccgtac accgaggtca ccataatgac actactcacc 420
241 agcatggccg ataaggagct ggtccacatg atcacctggg ccaagaagct tccaggtttc 480
242 ctgcagctgt ctctccatga ccaggtcgag ctgctggaga gctcgtggct ggaggtgctg 540
243 atgattgggc tcatatggag gtccatccac tgccccggca aactcatctt cgcacaggac 600
244 ctcatactgg acaggaatga aggtgactgt gtggaaggct ttgttgagat cttcgacatg 660
245 ctgctggcca ctgctcccg cttccgcatg ctcaaactca aacctgagga gtttgtctgc 720
246 ctcaaagcta tcactctgct caactctggt gccttctctt tctgcaccgg cacaatggag 780
247 cccctccaca acagcatggc agtgcagaac atgctggaca ccatcacaga cgctctcata 840
248 catcatatca gccaatcagg atgctcggct cagcagcagt cgaggcggca ggcccagctg 900
249 ctgctcctgc tctcccatc caggcacatg agcaacaaag gcatggagca tctctgcagc 960
250 atgaagtgca agaacaaagt gcctctgtac gacctg 996
252 <210> SEQ ID NO: 4
253 <211> LENGTH: 582
254 <212> TYPE: PRT

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/501,227

DATE: 08/18/2005

TIME: 16:46:40

Input Set : A:\600630-21US Seq List.txt

Output Set: N:\CRF4\08182005\J501227.raw